AD_____

Award Number: DAMD17-01-1-0421

TITLE: Prediction of Breast Cancer Risk by Aberrant Methylation in Mammary Duct Lavage

PRINCIPAL INVESTIGATOR: David M. Euhus, M.D., FACS

CONTRACTING ORGANIZATION: The University of Texas Southwestern Medical Center at Dallas
Dallas, TX 75390-9105

REPORT DATE: July 2005

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-07-2005	Annual	1 Jul 2004 - 30 Jun 2005

4. TITLE AND SUBTITLE
Prediction of Breast Cancer Risk by Aberrant Methylation in Mammary Duct Lavage

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-01-1-0421

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
David M. Euhus, M.D., FACS

E-Mail: david.euhus@utsouthwestern.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The University of Texas Southwestern Medical Center at Dallas
Dallas, TX 75390-9105

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Abstract follows.

15. SUBJECT TERMS
Breast cancer, aberrant methylation, methylation specific PCR, Nipple duct lavage risk assessment

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	20	

ABSTRACT

Women found to have atypical hyperplasia on a breast biopsy are at significantly increased risk for breast cancer. Nipple duct lavage (NDL) is being promoted by some as a new screening test for atypical hyperplasia, but cytological interpretation is a subjective art and experience indicates that the underlying conditions represented by cytological atypia on NDL can range from intraductal papilloma to ductal carcinoma in situ (DCIS). Laboratory studies indicate that methylation of tumor suppressor genes is an early event in breast carcinogenesis. We are applying objective methylation tests to cells obtained by NDL from women with breast cancer, women at increased risk of breast cancer, and women at low or average risk of breast cancer. Successful completion of the project will provide new tools for the objective evaluation of breast epithelial cells obtained by NDL in order to accurately risk stratify women and to enhance the early detection of DCIS. Thus far we have recruited and lavaged 148 women and have determined that atypia is diagnosed in only 25% of cancerous breasts. Atypia is diagnosed with equal frequency among fluid-producing and non-fluid producing ducts and the frequency of atypia declines with increasing age. We are currently running the methylation assays using a multiplexed quantitative methylation-specific real time PCR.

Table of Contents

INTRODUCTION

Women found to have atypical hyperplasia on a breast biopsy are at significantly increased risk for breast cancer. Nipple duct lavage (NDL) is being promoted by some as a new screening test for atypical hyperplasia, but cytological interpretation is a subjective art and experience indicates that the underlying conditions represented by cytological atypia on NDL can range from intraductal papilloma to ductal carcinoma in situ (DCIS). Laboratory studies indicate that methylation of tumor suppressor genes is an early event in breast carcinogenesis. We have previously shown that methylation of RASSF1A and APC in breast epithelial cells obtained by random fine needle aspiration biopsy correlates with breast cancer risk calculated using a validated mathematical model. In this same study we showed that methylation of Cyclin D2 is a breast cancer-specific change. Nipple duct lavage is a simple, minimally invasive approach for obtaining breast epithelial cells from most women, but the vagaries of cytological interpretation limit its usefulness for risk stratification or early diagnosis of cancer. Hence, we are evaluating tumor suppressor gene methylation in NDL samples as an adjunct or replacement for cytological assessment.

BODY

Recruitment and Sampling

Accrual and sampling are complete for all 150 patients. Required clinical data, including comprehensive breast cancer risk factor information, has been electronically archived for all of the patients. Cytological assessment has been completed for all of the samples and the results entered into the database. There were no serious adverse events from the NDL procedure.

Sample Analysis

As there is considerable uncertainty as to the optimal approach for extracting DNA from paucicellular clinical samples we performed a systematic comparison of five DNA extraction methods and have settled on the Puregene method for this study. We have nearly completed DNA extraction for all of the samples (mean of 3.0 ducts per patient) and as of this week have completed the bisulfite treatment for 160 samples.

Another DoD-funded investigator, Dr. Saraswati Sukumar, has recently described a multiplex quantitative methylation-specific real time PCR for the quantification of methylated gene copy numbers in clinical samples. We adopted the Sukumar method and optimized it in our laboratory this year. Of note, Mary Jo Fackler in Dr. Sukumar's laboratory. has been helpful in establishing data quality assurance criteria for our laboratory. We currently have the assay up and running for Cyclin D2, APC, RASSF1A, RAR-β, and HIN-1. Time and financial constraints dictate that this will be the final panel for the 150 patients included in this study.

We have completed the assays for four markers in 120 samples and are currently running assays all day every day.

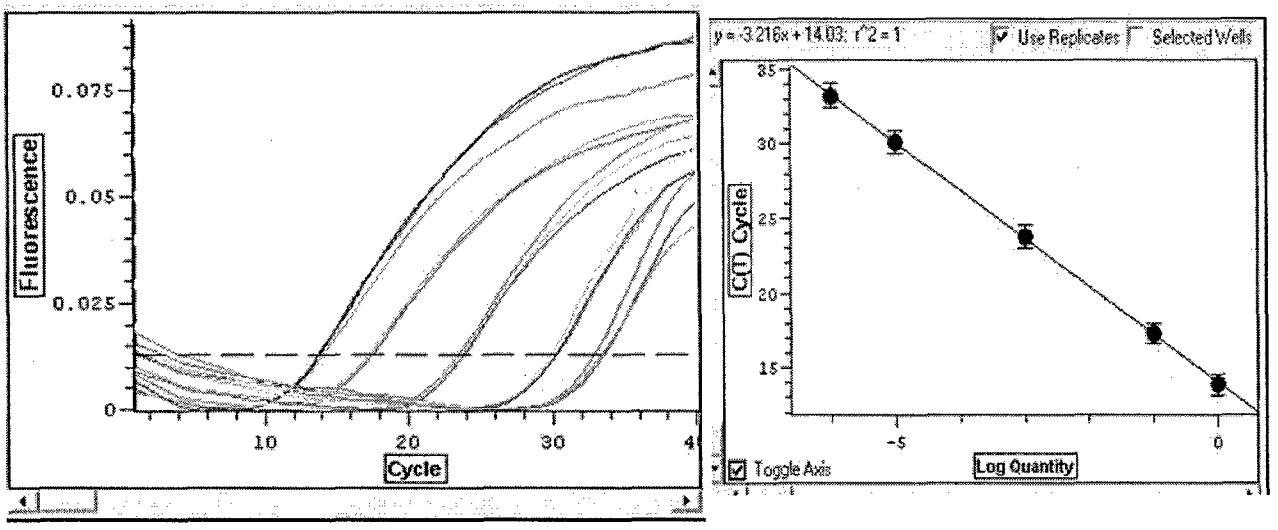

Figure: Sample Q-MS RTPCR output and a standard curve

Interim Analysis and Findings

Interim analysis has focused on quality of the samples and cytological findings. These results have been published in two papers which are included in the appendix so details will not be repeated here.

KEY RESEARCH ACCOMPLISHMENTS

- The prevalence of atypia is the same for fluid producing and non-fluid producing ducts (0.18, 95% CI 0.13 – 0.23 versus 0.15, 95% CI 0.09 – 0.20, P = NS).
- The prevalence of atypia declines with increasing age ($R^2 = 0.942$, P = 0.03).
- Atypia diagnosed by NDL has poor reproducibility (19% by duct, 34% by breast and 48% by patient)

REPORTABLE OUTCOMES

Abstract and presentation Society of Surgical Oncology, 2003[1]
Abstract and presentation Association for Academic Surgery, 2003[2]
Abstract and presentation Society of Surgical Oncology, 2004[3]
Abstract and presentation American Society of Breast Surgeons, 2004[4]
Abstract and Presentation 4th International Santa Barbara Symposium: The Intraductal Approach to Breast Cancer[5]
Manuscript published American Journal of Surgery, 2004[6]
Manuscript published, Cancer, 2005[7]

CONCLUSIONS

Preliminary analysis of our data highlights the limitations of cytological assessment as the sole modality for evaluation of nipple duct lavage samples. Preparatory work with methylation markers strongly suggests that the addition of molecular tests to the

evaluation of nipple duct lavage samples will provide an objective approach for risk stratification and perhaps for the early detection of mammographically occult DCIS.

REFERENCES

[1] Euhus DM, Ashfaq R, Milchgrub S, Naftalis E, Leitch AM, Virmani A. Comparison of Nipple Duct Lavage & Random Fine Needle Aspiration Biopsy for the Detection of Atypical Breast Epithelium. Soc Surg Oncol, Los Angeles, CA, March 2003.

[2] Cler LR, Ashfaq R, Naftalis E, Leitch AM, Hoover S, Euhus DM, Cytological Atypia Diagnosed by Nipple Duct Lavage: Reproducibility and MRI Findings, Assc Acad Surg, Sacramento, CA 2003.

[3] Johnson A, Ashfaq, R, Cler L, Naftalis E, Leitch AM, Hoover S, Euhus DM. Cytological Atypia Diagnosed by Nipple Duct Lavage: Reproducibility and MRI Findings. Soc Surg Oncol, New York, NY, March 2004.

[4] Johnson A, Ashfaq, R, Naftalis E, Leitch AM, Hoover S, Euhus DM. Patient and Duct Selection for Nipple Duct Lavage: Challenging the Existing Paradigm. American Society of Breast Surgeons, Las Vegas, NV, April 2004.

[5] Euhus D, Ashfaq R. Patient and Duct Characteristics Predicting Atypical Lavage Results, 4th International Santa Barbara Symposium: The Intraductal Approach to Breast Cancer, 2005

[6] Johnson A, Ashfaq, R, Naftalis E, Leitch AM, Hoover S, Euhus DM. Patient and Duct Selection for Nipple Duct Lavage. American Journal of Surgery, 2004;188:390-394.

[7] Johnson-Maddux A, Ashfaq R, Cler L, Naftalis E, Leitch A M, Hoover S, Euhus DM. Reproducibility of Cytologic Atypia in Repeat Nipple Duct Lavage. Cancer 2005;103:1129-1136.

Excerpta Medica

The American Journal of Surgery

The American Journal of Surgery 188 (2004) 390–394
Scientific paper

Patient and duct selection for nipple duct lavage

April Johnson Maddux, M.D.[a], Raheela Ashfaq, M.D.[b], Elizabeth Naftalis, M.D., F.A.C.S.[a], Ann Marilyn Leitch, M.D., F.A.C.S.[a], Susan Hoover, M.D., F.A.C.S.[a] David Euhus, M.D., F.A.C.S.[a,*]

[a]*Department of Surgery, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd., Dallas, TX 75390-9155, USA*
[b]*Department of Pathology, University of Texas Southwestern Medical Center at Dallas, Dallas, TX, USA*

Manuscript received May 19, 2004; revised manuscript June 6, 2004

Presented at the Fifth Annual Meeting of the American Society of Breast Surgeons, March 31–April 4, 2004, Las Vegas, Nevada

Abstract

Background: Nipple ductal lavage (NDL) is a new minimally invasive procedure with the potential to help identify women who could benefit from breast cancer risk intervention. NDL is currently encouraged for women with fluid-producing ducts and a 5-year Gail risk ≥1.7%. The purpose of this study was to evaluate the atypia rate by NDL in fluid-producing ducts compared with non–fluid-producing ducts and the atypia rate in high-risk verses low-risk patients to determine if current recommendations are supported.
Methods: Fifty-nine women were studied with NDL. The 226 ducts lavaged included all fluid-producing ducts (n = 136) and any dry ducts we could cannulate (n = 90). Breast cancer risk was calculated using mathematic models.
Results: There were 26 (44%) women with a 5-year Gail risk ≥1.7% and 33 (56%) with a 5-year Gail risk <1.7%. Cytologic atypia was diagnosed in 20 of 59 (34%) of patients. The atypia rate was similar for women with a 5-year Gail risk ≥1.7% (9 of 26 or 35%) compared with lower-risk women (11 of 33 or 33%, $P = 1.0$) and for fluid-producing ducts (26 of 136 or 19%) compared with dry ducts (14 of 90 or 15%, $P = 0.61$). No significant differences were found when the atypia was categorized as mild versus marked. Of note, the insufficient sample rate was higher for dry ducts (33%) compared with fluid-producing duct (22%, $P = 0.07$).
Conclusions: If NDL results are found to correlate with breast cancer incidence, it will be important to apply the test in a way that maximizes sensitivity for the detection of atypia in a screened population. We were unable to identify patient or duct characteristics that predict NDL atypia rates.

Keywords: Atypia; Cancer risk; Nipple duct lavage

The National Surgical Adjuvant Breast and Bowel Project (NSABP)-sponsored Breast Cancer Prevention Trial (BCPT-P1) reported that 5 years of tamoxifen therapy decreased the incidence of breast cancer by nearly 50% in women at increased risk for the disease [1]. For the purposes of this trial, increased risk was defined as a ≥1.7% probability of developing breast cancer during 5 years as calculated by the Gail model [2]. The ≥1.7% five-year Gail risk has been accepted by the Food and Drug Administration as providing a reasonable margin of benefit for tamoxifen therapy when balanced against the risks of thromboembolic complications and endometrial cancer. However, most women with a calculated 5-year breast cancer risk ≥1.7% refuse tamoxifen therapy. There is evidence, however, that women at higher risk levels, who are likely to enjoy greater benefits from tamoxifen, are more likely to accept this intervention. In this regard, women with previous breast biopsy specimens showing atypical ductal hyperplasia have a 3- to 5-fold greater risk of breast cancer than women with breast biopsy specimens showing nonproliferative fibrocystic changes only [3]. In addition, the NSABP BCPT-P1 trial recorded an 86% decrease in breast cancer incidence for these women. Based on these data, it is reasonable to direct chemoprevention resources toward women with atypical hyperplasia.

Previously, atypical ductal hyperplasia was only diagnosed incidentally when a palpable or mammographic abnormality was assessed by surgical biopsy. There has been

* Corresponding author. Tel.: +1-214-648-6467; fax: +1-214-648-7965
E-mail address: david.euhus@utsouthwestern.edu

0002-9610/04/$ – see front matter
doi:10.1016/j.amjsurg.2004.06.021

an interest, however, in developing more widely applicable, less-invasive approaches for identifying women with atypical breast epithelium. Wrensch et al [4,5] followed-up 2300 women for 12.7 years and reported that cytologic atypia in nipple aspirate fluid was associated with a relative risk for breast cancer of 4.9 and that the combination of cellular atypia with a family history of breast cancer was associated with a relative risk of 18. In the most recently reported follow-up, however, atypical cells in nipple aspirate fluid were associated with a relative risk for breast cancer of only 2.8 [6]. Fabian et al [7] found that high-risk women with atypical cells diagnosed by random fine-needle aspiration biopsy of the breast were 5 times more like to develop breast cancer than women without atypical cells.

Nipple duct lavage (NDL) has been proposed as a minimally invasive technique for obtaining breast epithelium for cytologic assessment. In a multicenter trial, Dooley et al [8] compared NDL with nipple duct aspiration (NDA) alone in 507 women and found that NDL was associated with a much lower insufficient sample rate than NDA (29% vs. 73%) because they retrieved an average of 13,500 epithelial cells compared with only 120 cells by NDA . Of note, the atypia rate for the increased risk women evaluated in this trial was 23% by NDL and 9% by NDA. It is currently unclear, however, whether atypia diagnosed by NDL confers the same risk of breast cancer as atypical hyperplasia diagnosed by surgical breast biopsy. Nevertheless, the test has been promoted as a method for breast cancer risk stratification and is currently recommended for women with a 5-year Gail risk $\geq$1.7% who would consider tamoxifen if they were found to have atypical cells. Because it is impractical to lavage each of the 6 to 16 duct orifices in each breast, it has been suggested that only fluid-producing ducts be lavaged. This study was designed to determine whether fluid-producing ducts are more likely to return cytologically atypical cells than dry ducts and to determine whether women with a 5-year Gail risk $\geq$1.7% are more likely to have atypia than lower-risk women.

Methods

Eligibility criteria

The Institutional Review Board at the University of Texas Southwestern Medical Center at Dallas approved this study, and written informed consent as well as Health Insurance Portability and Accountability Act authorizations were documented for each patient. Patients were recruited from the Mary L. Brown Breast Cancer Genetics and Risk Assessment Clinic in the University of Texas Southwestern Center for Breast Care. Comprehensive risk factor information was collected for each patient, and breast cancer risk calculated using the models of Gail, Claus, Bodian and BRCAPRO using software we developed (Breast C.A.R.E.). Certain components of this software are generally available in the CancerGene package we distribute [9]. All patients >18 years presenting for comprehensive breast cancer risk assessment were offered NDL on this protocol regardless of calculated risk level. Exclusion criteria included previous invasive breast cancer of any type; ductal carcinoma in situ or previous lobular carcinoma in situ treated by mastectomy; presence of a histologically undefined palpable or mammographic breast lesion suspicious for malignancy; bilateral prophylactic mastectomy; participation in a cancer prevention study (NSABP Protocol P-1 subjects who received placebo were eligible); any previous breast irradiation; any systemic chemotherapy in the past; performance status that restricted normal activity for a significant portion of each day; current use of androgens, luteinizing hormone–releasing hormone analogs, prolactin inhibitors, antiandrogens, or steroids (women who discontinue these drugs at least 3 months before duct lavage were eligible); any use of tamoxifen, raloxifene, or other selective estrogen-receptor modulator therapy; current use of coumadin; pregnant or lactating (within 6 month); presence of saline or silicone breast implants; or active bleeding disorder.

NDL procedure

EMLA cream (AstraZeneca, London, United Kingdom) was applied to the nipples and sealed with an occlusive plastic patch 2 hours before the procedure. The nipple area was dekeratinized by scrubbing with a mild abrasive gel. Breast massage was performed by the patient initially and then by the operator in an attempt to elicit nipple duct discharge. A nipple duct aspirator was used to encourage discharge; however, it was found that manual expression of fluid was generally more successful. An attempt was made to cannulate all fluid-producing ducts using a tapered dilator coated with 2% lidocaine jelly. If this was successful, a duct lavage catheter (Cytyk Health Corp., Boxborough, Massachusetts) was inserted, and a total of 10 mL physiologically buffered saline instilled and aspirated in 0.5-mL increments. When all fluid-producing ducts had been lavaged, an attempt was made to cannulate and lavage at least 1 non–fluid-producing duct in each breast.

Cytologic assessment

Lavage effluents were collected separately for each duct in 30 mL CytoLyt solution (Cytyk Health). Cytology slides were prepared using the thin-prep method, stained using the Papanicolaou technique, and then evaluated by a breast cytologist (R.A.). Cellularity was estimated for each sample as no cells or 1 to 10, 11 to 99, 100 to 999, or $\geq$1000 cells. A score (1 to 4) was assigned for each of 10 cytologic features: cellular arrangement, cell pleomorphism, myoepithelial cells, anisonucleosis, nucleoli, chromatin clumping, nuclear diameter, mitoses, nuclear molding, and cellular polarity. A composite score was calculated as the sum of the component scores. The cytologist was also asked to subjectively classify the sample according to all cell patterns

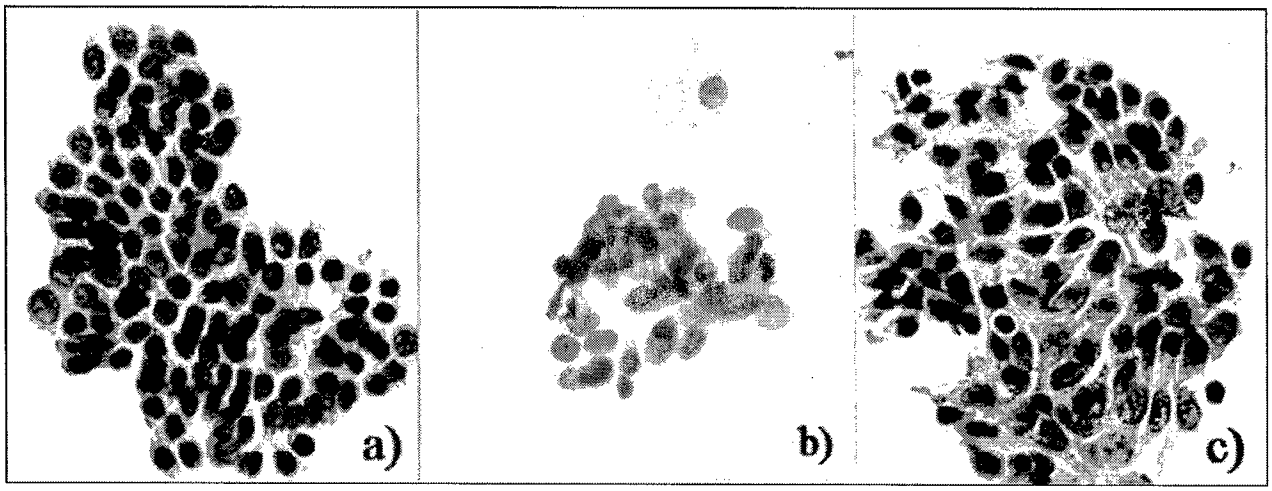

Fig. 1. Lavage cytology showing (**a**) normal cells, (**b**) mild atypia, and (**c**) marked atypia.

observed: normal, apocrine metaplasia, typical hyperplasia, mild atypia, marked atypia, or suspicious for cancer. Examples of normal cells, cells with mild atypia, and cells with marked atypia are shown in Fig. 1.

Results

Patients

NDL was performed for 59 asymptomatic women whose risk of breast cancer had been defined using mathematic models. Most of the patients were white (95%), and most were premenopausal (56%). The median age was 43 years. Twenty-six (44%) had a 5-year Gail risk ≥1.7%, and 33 (56%) had a 5-year Gail risk <1.7% (Table 1). A total of 226 ducts were lavaged. Of these, 136 were fluid producing, and 90 were dry. The insufficient sample rate was 30 of 136 (22.1%) for the fluid-producing ducts and 30 of 90 (33.3%) for the dry ducts ($P = 0.07$).

Table 1
Patient demographics

No. of patients	59
No. age in years (%)	
21–30	4 (7)
31–40	19 (32)
41–50	25 (42)
51–60	8 (14)
61–70	3 (5)
No. race (%)	
White	56 (95)
Hispanic	2 (3)
Asian	1 (2)
African American	0
No. menopausal status (%)	
Premenopausal	35 (59)
Perimenopausal	6 (10)
Postmenopausal	18 (31)
No. 5-year Gail risk	
≥1.7%	26 (44)
<1.7%	33 (56)

Atypia rate by Gail risk calculation

Overall, atypia was diagnosed in 20 of 59 (34%), mild atypia in 13 of 59 (22%), and marked atypia in 7 of 59 (12%) patients. The atypia rate was similar for women with a 5-year Gail risk ≥1.7% compared with women having a 5-year Gail risk <1.7% (9 of 26 or 35% vs. 11 of 33 or 33%, respectively, $P = 1.0$). Marked atypia was more common in women with a 5-year Gail risk ≥1.7% (4 of 26, 15%) than in women with a 5-year Gail risk <1.7% (3 of 33 9%), but this result was not statistically significantly ($P = 0.73$, Table 2).

Atypia rate by fluid-producing status of ducts

Overall, atypia was diagnosed in 40 of 226 (18%) ducts, mild atypia in 28 of 226 (12%), and marked atypia in 12 of 226 (5%). The atypia rate was similar for fluid-producing (26 of 136 or 19%) and dry ducts (14 of 90 or 15%, $P = 0.61$) with no differences noted for mild or marked atypia (Table 3).

Atypia rate by fluid-producing status and calculated Gail risk

It is currently recommended that only patients with a 5-year Gail risk ≥1.7% and fluid-producing ducts undergo

Table 2
Atypia rate by calculated Gail risk

Gail risk	All atypia	Mild atypia	Marked atypia
No. all patients (%)	20/59 (34)	13/59 (22)	7/59 (12)
No. 5-year Gail risk (%)			
≥1.7%	9/26 (35)*	5/26 (19)†	4/26 (15)‡
<1.7%	11/33 (33)*	8/33 (24)†	3/33 (9)‡

* $P = 1.00$; † $P = 0.89$; ‡ $P = 0.73$.

Table 3
Atypia rate by fluid-producing status of ducts

Duct status	Any atypia	Mild atypia	Marked atypia
No. all ducts (%)	40/226 (18)	28/226 (12)	12/226 (5)
No. fluid producing (%)	26/136 (19)*	18/136 (13)†	8/136 (6)‡
No. dry ducts (%)	14/90 (15)*	10/90 (11)†	4/90 (4)‡

* $P = 0.61$; †$P = 0.80$; ‡$P = 0.88$.

ductal lavage for additional risk stratification. We calculated atypia rates for patients with a 5-year Gail risk ≥1.7% and <1.7% considering results only for fluid-producing or only for non–fluid-producing ducts. The atypia rate for patients with a 5-year Gail risk ≥1.7% considering only fluid-producing ducts was 7 of 22 (32%). The atypia rate for patients with a 5-year Gail risk <1.7% considering only non–fluid-producing ducts was 8 of 27 (30%, $P = 1.00$). Of note, although atypia rates for dry and fluid-producing ducts among patients with a 5-year Gail risk <1.7% were similar (25% vs. 30%, $P = 0.93$), the atypia rate was higher for fluid-producing than dry ducts for patients with a 5-year Gail risk ≥1.7% (32% vs. 11%, Fig. 2), but this result did not reach statistical significance ($P = 0.09$).

Comments

NDL is currently proposed as a minimally invasive approach for identifying atypical breast epithelial cells for the purpose of individualized breast cancer risk stratification. If NDL results are shown to correlate with breast cancer incidence, it will be important to apply the test in a way that maximizes its sensitivity for detection of atypical cells in the screened population. Current recommendations, however, limit the test to women with fluid-producing ducts and a 5-year Gail risk ≥1.7%. Both of these criteria are reasonably challenged based on previously published studies.

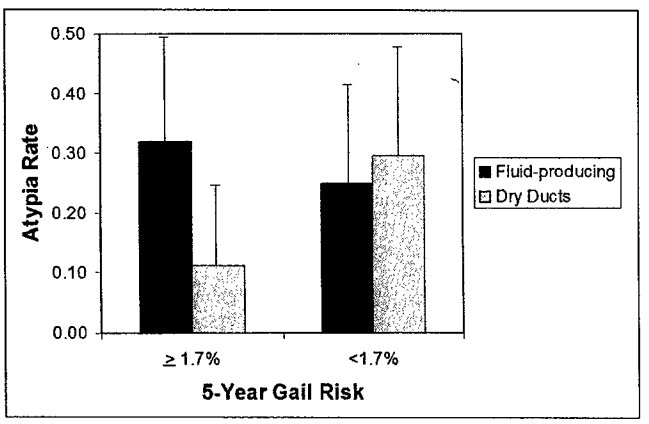

Fig. 2. Atypia rates by calculated Gail risk and fluid-producing status of the ducts. Atypia rates were similar for all categories except for fluid-producing versus dry ducts among women with a calculated 5-year Gail risk ≥1.7% (32% vs. 11%, $P = 0.09$).

First, data from the Nurses Health Study demonstrated that during a 5-year period, 753 of 54,844 women with a 5-year Gail risk <1.7% developed breast cancer compared with 601 of 27,225 women with a 5-year Gail risk ≥1.7% [10]. That is, most of the breast cancers observed in this cohort (55%) occurred in women who would not have been considered eligible for ductal lavage. Second, the etiology of expressible nipple duct discharge is multifactorial (e.g., duct ectasia, apocrine metaplasia, papilloma), and the presence of expressible nipple duct fluid does not reliably distinguish patients with benign breast disease from control patients [11]. Based on these observations, it is reasonable to ask whether the atypia rate, as measured by NDL, is higher for patients with 5-year Gail risks ≥1.7% than for lower risk women or for fluid-producing ducts compared with dry ducts. Our data suggested that these criteria do not identify women that are more likely to have atypical cells diagnosed by NDL.

The most obvious limitation of this study was the small sample size and the possibility of a type II error. It should be noted, however, that with respect to atypia rates in fluid-producing versus dry ducts, with α set at 0.05, our study had a power of 0.979 to recognize a 20% difference (30% vs. 10%). With respect to atypia rates in women with a 5-year Gail risk ≥1.7% versus <1.7%, our power was only 0.486 to recognize a similar difference. Even if a larger study were to demonstrate a statistically significant difference between atypia rates in high- and low-risk women, it is unlikely that the difference would be clinically significant in the context of population screening.

The difference in atypia rates between fluid-producing and dry ducts among the increased risk women (Fig. 2) is intriguing although not statistically significant. It is possible that there are important biologic differences between the epithelial cells retrieved from women at different risk levels that cannot be recognized under the microscope. If this is the case, limiting NDL to women determined to be at increased risk based on epidemiologic models would be reasonable. This can only be known, however, as the results of several on going studies evaluating biomarker expression in lavage cells become available.

NDL is an intriguing technology for sampling breast epithelial cells from selected nipple ducts. It remains to be determined, however, how best to select the duct(s) to lavage, how best to select the patients to lavage, and, most important, whether atypia diagnosed by NDL predicts an increased risk for breast cancer. Clearly, additional study is required.

References

[1] Fisher B, Constantino JP, Wickerham DL, et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. J Natl Cancer Inst 1998;90: 1371–88.

[2] Gail MH, Brinton LA, Byar DP, et al. Projecting individualized

probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 1989;81:1879–86.
[3] Dupont WD, Parl FF, Hartman WH, et al. Breast Cancer risk associated with proliferative breast disease and atypical hyperplasia. Cancer 1993;71:1258–65.
[4] Wrensch MR, Petrakis NL, King EB, et al. Breast cancer incidence in women with abnormal cytology in nipple aspirates of breast fluid. Am J Epidemiol 1992;135:130–41.
[5] Wrensch MR, Petrakis NL, King EB, et al. Breast cancer risk associated with abnormal cytology in nipple aspirates of breast fluid and prior history of breast biopsy. Am J Epidemiol 1993;137:829–33.
[6] Wrensch MR, Petrakis NL, Miike R, et al. Breast cancer risk in women with abnormal cytology in nipple aspirates of breast fluid. J Natl Cancer Inst 2001;93:1791–8.
[7] Fabian CJ, Kimler BF, Zalles CM, et al. Short-term breast cancer prediction by random periareolar fine-needle aspiration cytology and the Gail risk model. J Natl Cancer Inst 2000;92:1217–27.
[8] Dooley WC, Ljung B, Veronesi U, et al. Ductal lavage for the detection of cellular atypia in women at high risk for breast cancer. J Natl Cancer Inst 2001;93:1624–32.
[9] Euhus DM. Understanding mathematical models for breast cancer risk assessment and counseling. Breast J 2001;7:224–32.
[10] Rockhill B, Spiegelman D, Byrne C, et al. Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. J Natl Cancer Inst 2001;93:358–66.
[11] Wynder EL, Lahti H, Laakso K, et al. Nipple aspirates of breast fluid and the epidemiology of breast diseases. Cancer 1985;56:1473–8.

Reproducibility of Cytologic Atypia in Repeat Nipple Duct Lavage

April Johnson-Maddux, M.D.[1]
Raheela Ashfaq, M.D.[2]
Leslie Cler, B.S.[1]
Elizabeth Naftalis, M.D.[1]
Ann Marilyn Leitch, M.D.[1]
Susan Hoover, M.D.[1]
David M. Euhus, M.D.[1]

[1] Department of Surgery, University of Texas Southwestern Medical Center, Dallas, Texas.

[2] Department of Pathology, University of Texas Southwestern Medical Center, Dallas, Texas.

Supported by a grant from the Department of Defense Breast Cancer Research Program (DAMD17-01-1-0421).

Address for reprints: David M. Euhus, M.D., Division of Surgical Oncology, E6.222, 5323 Harry Hines Boulevard, Dallas, TX 75390-9155; Fax: (214) 648-7965; E-mail: david.euhus@utsouthwestern.edu

Received August 4, 2004; revision received October 19, 2004; accepted October 28, 2004.

BACKGROUND. It is believed that atypical cells identified by nipple duct lavage (NDL) indicate an increased risk for breast carcinoma similar to atypical ductal hyperplasia diagnosed by tissue biopsy, but many basic performance characteristics of NDL currently are undefined.

METHODS. NDL was performed in 108 patients unselected for breast carcinoma risk and then was repeated after 2–14 months (median, 8 months) if the initial lavage was classified as atypical. Breast magnetic resonance images (MRIs) were obtained from a subset of patients who had atypical lavage results.

RESULTS. Marked atypia was diagnosed in 22% of 36 breasts with an incident carcinoma compared with 7% of 172 unaffected breasts ($P = 0.01$). After excluding breasts with an incident carcinoma, there were 32 patients (30%) with either mild or marked atypia. The lavage was repeated in 23 of these women, and the second lavage was classified as atypical in 48%. Neither marked atypia on the initial lavage nor a 5-year Gail risk $\geq$ 1.7% predicted atypia on repeat lavage, but there was a trend for improved reproducibility when the atypia initially was diagnosed in a fluid-producing duct. MRIs were abnormal in 13% of 24 breasts with an atypical lavage, and ductal carcinoma in situ was diagnosed subsequently in 1 breast.

CONCLUSIONS. Atypia frequently is diagnosed by NDL, but the reproducibility of repeat lavage is low. Lavage atypia may be physiologic or artifactual rather than pathologic in many instances. Marked atypia occasionally may represent mammographically occult ductal carcinoma in situ. ***Cancer* 2005;103:1129–36.**

KEYWORDS: breast neoplasms, precancerous conditions, nipples, epithelial cells.

Most breast carcinomas arise from the epithelial cells lining the ductal system, and atypical ductal epithelium is a marker of increased risk for the development of breast carcinoma. Wresch et al. followed 2300 women over 12.7 years and found that cytologic atypia in nipple aspirate fluid (NAF) was associated with a relative risk for breast carcinoma of 4.9.[1] An updated analysis of those data adjusted the relative risk down to 2.8.[2] Fabian et al. reported that women with a personal history of invasive or in situ breast carcinoma or with a 10-year Gail breast carcinoma risk > 4% who were found to have atypical cells on random fine-needle aspiration breast biopsy (FNAB) were 5 times more likely to develop breast carcinoma than women with a 10-year Gail risk < 4% and no atypia.[3] Dupont et al. found that the risk of breast carcinoma was increased 4.3-fold in women who were diagnosed with atypical hyperplasia by surgical biopsy.[4] Nipple duct lavage (NDL) has been proposed as a minimally invasive technique for obtaining breast epithelial cells, with the assumption that atypia identified by NDL confers the same relative risk for breast carcinoma as atypia identified by NAF, FNAB, or surgical biopsy.

DOI 10.1002/cncr.20884
Published online 31 January 2005 in Wiley InterScience (www.interscience.wiley.com).

Although this has not been confirmed in a prospective trial to date, there is justifiable interest in developing and validating a minimally invasive procedure for the detection of atypical hyperplasia.

The National Surgical Adjuvant Breast Project (NSABP)-sponsored Breast Cancer Prevention Trial (BCPT-P1) demonstrated that 5 years of tamoxifen reduced the risk of breast carcinoma by $\approx$ 50% in increased risk women,[5] but most eligible women refuse to take tamoxifen.[6] Because women with atypical hyperplasia are at significantly increased risk for breast carcinoma and experience the greatest risk reduction with tamoxifen (86%), a test for atypia, such as NDL, may help eligible women decide to accept tamoxifen.

Although clinical guidelines for NDL were published previously.[7,8] many of the basic performance characteristics of the procedure remain unknown. A multiinstitutional study comparing NDL with NAF found that the insufficient sample rate was much lower for NDL than for NAF (22% vs. 73%) and that the atypia rate was much higher (24% vs. 10%).[9] However, currently, it is unknown whether atypia diagnosed by NDL predicts an increased risk for breast carcinoma. Some atypical lavages may reflect underlying atypical hyperplasia, but others are likely to reflect reversible physiologic changes related to the hormonal milieu, benign intermediate-risk lesions (such as intraductal papilloma or papillomatosis), or fully developed ductal carcinoma in situ (DCIS). Lavage atypia that is not reproducible may be related to reversible physiologic changes in the breast epithelium, whereas atypia that is reproduced may be related to fixed, underlying pathologic alterations. We performed repeat NDL and magnetic resonance imaging (MRI) scans in women with lavage atypia to estimate the prevalence of persistent lavage atypia and the prevalence of mammographically occult DCIS or invasive carcinoma when atypical cells are identified.

MATERIALS AND METHODS

Eligibility Criteria

The Institutional Review Board at the University of Texas Southwestern Medical Center at Dallas approved the protocol for this study, and written informed consent was documented for all participants. Patients from the Mary L. Brown Cancer Genetics and Risk Assessment Clinic at the University of Texas Southwestern Center for Breast Care were recruited for enrollment. A comprehensive risk assessment was performed for each patient. Patients with incident breast carcinoma and unaffected women age $\geq$ 18 years who presented for breast carcinoma risk assessment were offered duct lavage regardless of their calculated risk level. Exclusion criteria included the presence of an undefined palpable or mammographic breast lesion suspicious for malignancy; bilateral prophylactic mastectomy; any prior breast irradiation; any systemic chemotherapy in the past; a performance status that restricted normal activity for a significant portion of the day; current use of androgens, luteinizing hormone-releasing hormone analogs, prolactin inhibitors, antiandrogens, or corticosteroids (women were eligible if these drugs were discontinued 3 months prior to lavage); ever use of tamoxifen, raloxifene, or other selective estrogen receptor modulator (SERM) therapy; or pregnancy or lactation within 6 months.

NDL Procedure

Local anesthetic cream (EMLA; AstraZeneca, London, United Kingdom) was applied to the nipple and then covered with an occlusive patch 1–2 hours prior to the procedure. At the start of the procedure, the patient performed a self-breast massage, after which, the nipple was dekeratinized with a mild abrasive gel (Nuprep; D. O. Weaver and Company, Aurora, CO). The operator then continued the breast massage in an effort to express NAF. If no NAF was elicited manually, then a nipple aspirator (FirstCyte; Cytyc Health Corporation) was used. Fluid-producing ducts initially were cannulated with a tapered dilator coated with 2% lidocaine jelly, after which, a ductal lavage microcatheter (FirstCyte Microcatheter; Cytyc Health Corporation) was inserted. Saline (10 mL) was infused into the duct in 0.5-mL increments, and the effluent fluid was aspirated. An attempt was made to lavage all fluid-producing ducts and at least one nonfluid-producing duct from each breast. The location of each cannulated duct orifice was recorded on a circular grid with 45 cells, so that the orifice of any duct that yielded atypical cells could be identified in the future. Repeat lavage was offered to women whose initial lavage returned atypical cells. At the time of repeat lavage, every effort was made to recannulate the same ducts that were cannulated at the initial lavage. Breast MRI was recommended for all women whose initial lavage returned atypical cells but was performed only if third-party payor approval could be obtained.

Cytologic Evaluation

The lavage effluents from each duct were collected separately in 30 mL of CytoLyt solution (Cytyc Health Corporation). Cytology slides were prepared using the ThinPrep method and were stained using the Papanicolaou technique. All slides were evaluated by the same breast cytopathologist (R.A.), who classified each sample according to the most severe alterations iden-

tified: insufficient for diagnosis, normal epithelium or apocrine metaplasia only, typical epithelial hyperplasia, mild atypia, or marked atypia. Cytologic interpretation was performed according to the guidelines published by the Cytyc Health Corporation (http://www.ductallavage.com/professionals/cytologyTraining.cfm). Briefly, mild atypia was defined as clusters of crowded, overlapping cells with slight nuclear enlargement, mild anisonucleosis, prominent nucleoli, occasional myoepithelial cells, and granular, evenly distributed chromatin. Marked atypia was diagnosed when these same features were more pronounced and included marked anisonucleosis, significantly increased nuclear:cytoplasmic ratios, and irregular, clumping chromatin.

Statistical Analysis

Proportions and atypia prevalence rates were compared using the Fisher exact test. The α value was set at 0.05.

RESULTS

Demographics

Ductal lavage was performed in 377 ducts from 208 breasts in 108 female patients. On average, 1.8 ducts were lavaged per breast, and 3.5 ducts were lavaged per patient. There were 41 women with incident breast carcinoma and 67 women who were unaffected with breast carcinoma but who had completed a comprehensive breast carcinoma risk assessment. Among the women who were unaffected with breast carcinoma, 52% had a 5-year Gail risk $< 1.7\%$, and 48% had a 5-year Gail risk $\geq 1.7\%$. The mean patient age was 46.3 years (range, 30–82 years), and 42% of patients were postmenopausal. Most of the patents were Caucasian (81%), and 35% of patients were using oral contraceptives or hormone replacement therapy at the time of initial sampling. NAF was expressible from 86% of the patients (Table 1).

TABLE 1
Characteristics of the Study Sample

Characteristic	No. of patients (%)
Total patients	108 (100.0)
Age (yrs)	
Mean	46.3
Range	30.0–81.5
Ethnicity	
Caucasian	87 (80.6)
African American	16 (14.8)
Hispanic	4 (3.7)
Asian	1 (0.9)
Expressible nipple aspirate fluid	93 (86.1)
Menopausal status	
Premenopausal	56 (51.9)
Perimenopausal	7 (6.5)
Postmenopausal	45 (41.7)
Oral contraceptive use (premenopausal)	16/56 (28.6)
Hormone replacement (perimenopausal and postmenopausal)	22/52 (42.3)
Risk groups	
Breasts ipsilateral to a breast carcinoma	36 (100.0)
DCIS only	3/36 (8.3)
Infiltrating ductal carcinoma	29/36 (80.6)
Infiltrating lobular carcinoma	3/36 (8.3)
Medullary carcinoma	1/36 (2.8)
Any associated DCIS	29/36 (80.6)
Breasts contralateral to a breast carcinoma	38
Unaffected risk assessed patients	67/108 (62.0)
History of ADH	1/67 (1.5)
BRCA gene mutation	3/67 (4.5)
5-Yr Gail risk	
0.01–0.85	20/67 (29.9)
0.86–1.69	15/67 (22.4)
1.70–2.54	17/67 (25.4)
> 2.54	15/67 (22.4)

DCIS: ductal carcinoma in situ; ADH: atypical ductal hyperplasia.

Frequency of Atypia

Table 2 summarizes the atypia rates for 36 breasts with an incident breast carcinoma, 38 breasts contralateral to an incident breast carcinoma, and 134 breasts from women who were unaffected by breast carcinoma. Results are reported separately for the right and left breasts of the unaffected women to permit comparisons between patients with breast carcinoma and unaffected patients on a per-breast basis. The insufficient sample rate was higher for ducts from breasts with an incident carcinoma (40%) than for ducts from breasts that were unaffected with breast carcinoma (27%; $P = 0.06$). Atypia of any degree was diagnosed in 36% of breasts with an incident breast carcinoma and in 24% of breasts that were unaffected with breast carcinoma ($P = 0.19$), but marked atypia was diagnosed more frequently in breasts with an incident breast carcinoma (22%) than in unaffected breasts (7%; $P = 0.01$). Among breasts that were unaffected with breast carcinoma, we diagnosed cytologic atypia in 18% of ducts, 24% of breasts, and 30% of patients. There were no trends in the insufficient sample rate or in the frequency of diagnosis of mild or marked atypia for initial lavages over time (Fig. 1).

Reproducibility of Atypia

Among the 32 patients who had an atypical lavage from a breast that was unaffected with breast carcinoma, repeat lavage was performed for 23 patients. Four patients with breast carcinoma received chemotherapy after the initial lavage, which rendered them ineligible for repeat lavage of the contralateral breast; two women without breast carcinoma moved out of

TABLE 2
Frequency of Atypia by Sampling Group

	No. of patients (%)							
	Ducts				Breasts			
Variable	ICMD	Mild atypia	Marked atypia	Any atypia	ICMD	Mild atypia	Marked atypia	Any atypia
Cancerous breast	24/60 (40.0)	7/60 (11.7)	8/60 (13.3)	15/60 (25.0)	11/36 (30.6)	5/36 (13.9)	8/36 (22.2)	13/36 (36.1)
Contralateral breast	19/57 (33.3)	6/57 (10.5)	3/57 (5.3)	9/57 (15.8)	10/38 (26.3)	5/38 (13.2)	3/38 (7.9)	8/38 (21.1)
Unaffected right breast	35/133 (26.3)	15/133 (11.3)	7/133 (5.3)	22/133 (16.5)	10/67 (14.9)	12/67 (17.9)	5/67 (7.5)	17/67 (25.4)
Unaffected left breast	32/127 (25.2)	19/127 (15.0)	6/127 (4.7)	25/127 (21.3)	13/67 (19.4)	12/67 (17.9)	4/67 (6.0)	16/67 (23.9)

ICM: insufficient cellular material for diagnosis.

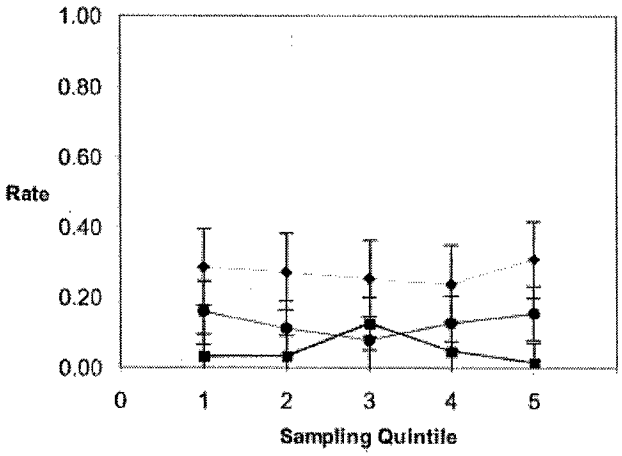

FIGURE 1. Insufficient cellular material for diagnosis (ICMD) rate and frequency of mild and marked atypia for initial lavages over time. The entire study sample was divided into quintiles to detect trends in the insufficient sample rate or in the frequency of diagnosis of mild or marked atypia that might suggest systematic changes in the performance or interpretation of nipple duct lavage over time. No such trends were identified. Error bars bracket the 95% confidence intervals. ◆: ICMD; ●: Mild atypia; ■: Marked atypia.

the area; two women declined repeat lavage; and one woman underwent mastectomy without repeating the lavage after an MRI was interpreted as highly suspicious (Fig. 2). Repeat lavage was performed 2.3–14.3 months (median, 8.3 months) after the initial lavage. Because every effort was made to relavage all of the ducts that had been lavaged initially in the patients who had at least 1 duct diagnosed as atypical, a total of 78 ducts from 45 breasts were relavaged in these 23 patients. If any duct was classified as atypical on the repeat lavage, then the atypia was scored as "reproducible" for that patient. The repeat lavage was classified as atypical for 11 of 23 patients (48%), 11 of 32 breasts (34%), and 8 of 42 ducts (19%) that were diagnosed initially as atypical (Table 3). Failure to reproduce the atypia was due to an insufficient sample on the second lavage in 13% of patients and due to a diagnosis of normal epithelium or typical hyperplasia only in 39% of patients. Marked atypia on the initial lavage was no more predictive of an atypical second lavage than mild atypia (44% vs. 50%; $P = 1.0$).

Among the patients who produced NAF, atypia was diagnosed on the second lavage in 55%, compared with 0% for the 3 patients who did not produce NAF ($P = 0.25$). Among the patients who had a 5-year Gail risk $\geq 1.7\%$, atypia was diagnosed on the second lavage in 22% of patients, compared with 70% of the patients who had a 5-year Gail risk $< 1.7\%$ ($P = 0.10$). Reproducibility rates were similar for premenopausal women compared with perimenopausal or postmenopausal women, for women who were taking hormonal medications compared with women who were not taking these medications, and for women who underwent repeat lavage < 8.3 months after the initial lavage (the median interval for this series) compared with women who underwent repeat lavage ≥ 8.3 months after the initial lavage.

MRI Findings

MRI was performed in 24 breasts from 17 women whose initial lavage was interpreted as atypical. The atypia was marked in 9 breasts, and the MRI was abnormal in 1 of those breasts (11%). Total mastectomy revealed 10 cm of DCIS in this patient (Fig. 2). Repeat lavage was not performed prior to the MRI and subsequent surgery. The initial lavage was interpreted as mildly atypical in 15 breasts; and of those, the MRI was abnormal in 2 breasts (13%). The MRI was interpreted as borderline suspicious in both of those breasts. Repeat MRI in one patient demonstrated resolution of the region of abnormal enhancement, and repeat lavage in this patient was classified as typical epithelial hyperplasia only. In the second patient with

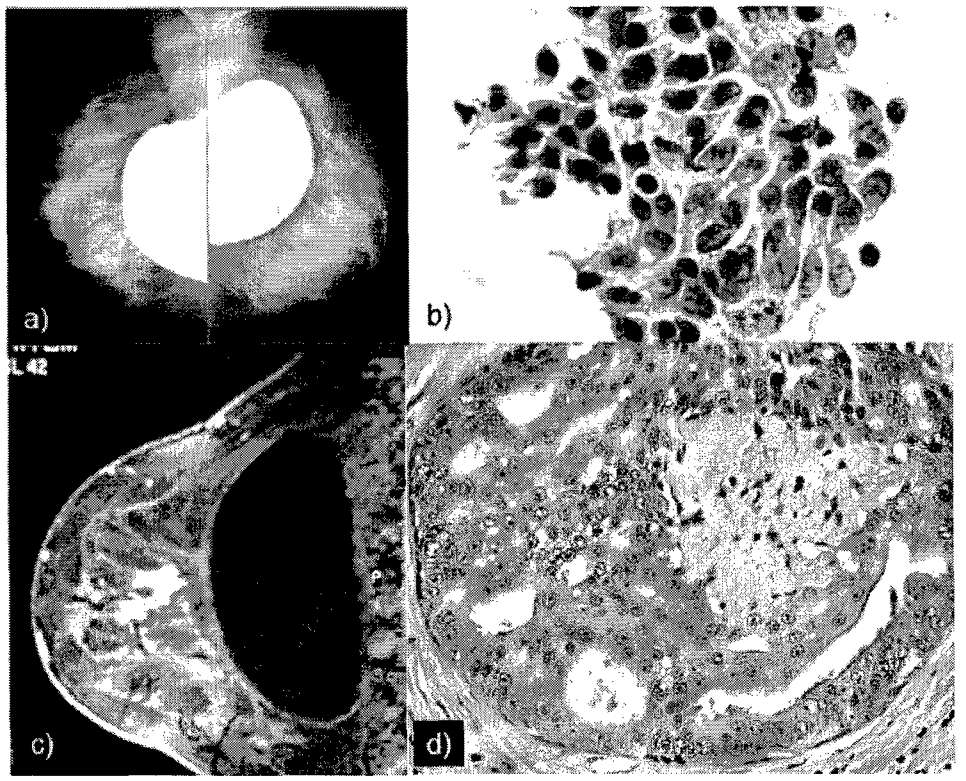

FIGURE 2. An interesting patient. (a) A screening mammogram revealed a suspicious cluster of microcalcifications in the right breast. The left breast was interpreted as normal. (b) Nipple duct lavage of the left breast at the time of right mastectomy returned markedly atypical cells. (c) This left breast magnetic resonance image demonstrates three areas of abnormal enhancement. (d) A subsequent left total mastectomy revealed 10 cm of ductal carcinoma in situ.

TABLE 3
Classification of Repeat Lavage Results According to Initial Lavage Results

	Initial lavage: No. of patients (%)										
	Mild atypia					Marked atypia					
Repeat lavage	Marked	Mild	Normal or EH	ICMD	Any atypia	Marked	Mild	Normal or EH	ICMD	Any atypia	Any atypia
By patient	3/14 (0.21)	4/14 (0.29)	5/14 (0.36)	2/14 (0.14)	7/14 (0.50)	2/9 (0.22)	2/9 (0.22)	4/9 (0.44)	1/9 (0.11)	4/9 (0.44)	11/23 (0.48)
By breast	4/22 (0.18)	5/22 (0.23)	9/22 (0.41)	4/22 (0.18)	9/22 (0.41)	1/10 (0.10)	1/10 (0.10)	6/10 (0.60)	2/10 (0.20)	2/10 (0.20)	11/32 (0.34)
By duct	3/30 (0.10)	4/30 (0.13)	13/30 (0.43)	10/30 (0.33)	7/30 (0.23)	0/12 (0.00)	1/12 (0.08)	9/12 (0.75)	2/12 (0.17)	1/12 (0.08)	8/42 (0.19)

EH: typical epithelial hyperplasia; ICMD: insufficient cellular material for diagnosis.

mild atypia, a repeat lavage was classified as marked atypia, and a targeted ultrasound examination was unremarkable. Repeat MRI in this patient was interpreted as entirely normal.

DISCUSSION

NDL has been proposed as a secondary risk-stratification procedure for women who are determined to be at increased risk for breast carcinoma based on epidemiologic models. It is believed that the atypical cells identified by NDL confer the same breast carcinoma risk as atypia found in nipple aspirates, needle biopsies, and surgical biopsies; however, currently, there are no data to confirm this, and many of the essential performance characteristics of NDL are yet to be elucidated.

Although NDL currently is recommended for risk stratification, and not for early detection of breast carcinoma, it is reasonable to suppose that, if lavage atypia is a strong predictor of breast carcinoma risk,

TABLE 4
Comparison of Insufficient Cellular Material for Diagnosis and Atypia Rates in the Current Series with Previously Reported Rates

	No. of patients (%)	
Results	Dooley et al., 2001[9]	Current series
By patient		
ICMD	84/383 (22)	17/105 (16)
Mild atypia	66/383 (17)	19/105 (18)
Marked atypia	26/383 (7)	13/105 (12)
By duct		
ICMD	173/591 (29)	86/317 (27)
Mild atypia	77/591 (13)	40/317 (13)
Marked atypia	28/591 (5)	16/317 (5)

ICMD: insufficient cellular material for diagnosis.

then it would occur at a high frequency in breasts with an incident breast carcinoma. Marked atypia, in fact, was more common in breasts with an incident breast carcinoma (22%) than in unaffected breasts (6–8%). This is a higher marked atypia rate than that reported for a series of 28 mastectomy patients.[10] In that study, atypia rates were reported per lavage sample (ducts) rather than by breast, and marked atypia was identified in only 4 of 29 (14%) adequate samples from patients with an incident breast carcinoma. The exclusion of patients who had insufficient samples from the current series resulted in a marked atypia rate of 8 of 25 (32%) for breasts with an incident carcinoma compared with 12 of 139 (9%) for unaffected breasts ($P = 0.003$). In the prior study, the extent of carcinoma in situ appeared to correlate with the degree of atypia identified in the lavage samples, but DCIS was identified in 23 of 28 patients (82%), and it was not possible to determine whether or not lavage atypia was associated primarily with DCIS. Similarly, 29 of 36 patients (81%) with incident breast carcinoma in the current series had a DCIS component, but marked atypia was identified in 3 of 7 patients (42%) who had no DCIS component. It is clear that the marked atypia rate for breasts with an incident breast carcinoma is considerably higher than that previously reported and is considerably higher than that observed for breasts that are unaffected by breast carcinoma.

The prevalence of lavage atypia measured in our series of patients, who were unselected for breast carcinoma risk, was nearly identical to that reported in the first large validation series, a series that was limited to women who were at increased risk for breast carcinoma[9] (Table 4). To make this comparison, we have considered only breasts that were unaffected by breast carcinoma (38 breasts contralateral to a breast carcinoma and 134 breasts from women who were unaffected with breast carcinoma). Results for 3 women with breast carcinoma who did not undergo lavage of the contralateral breast were excluded, leaving a total of 105 patients for comparison.

Atypia was reproduced on repeat lavage in only 8 of 42 ducts (19%), 11 of 32 breasts (34%), and 11 of 23 patients (48%). A recent series of duct lavages in 38 high-risk women reported an atypia prevalence of 23% for those with an adequate sample and reproducibility of the atypia in only 1 of the 4 patients who underwent repeat lavage.[11] The reproducibility of any cytology-based screening test will be related to the physiologic factors that affect the cytologic features of the cells collected, variability in the sampling procedure, and variability in the cytologic interpretation. The breast is an exquisitely hormone-responsive organ, and fluctuations in the hormonal milieu may affect cytology-based screening tests. Exogenous estrogens have been associated with atypical hyperplasia in humans,[12] monkeys,[13] and rodents,[14] but it is not known whether these lesions are reversible in humans when the hormones are withdrawn. It is noteworthy that nine of our patients who underwent repeat lavage were using hormonal medications at the time of the initial lavage. Atypia was reproduced in 2 of the 4 patients (50%) who discontinued these medications between the first and second lavages. Endogenous estrogens also may influence the cytologic appearance of breast epithelial cells, although Mitchell et al. found no significant changes in breast epithelial cells recovered from weekly NAF samples that were collected over two menstrual cycles.[15]

Although every effort was made to recannulate the same ducts that were diagnosed as atypical on the initial lavage, technical problems with resampling may have contributed to the low reproducibility. The shear numbers of duct orifices clustered near the center of the papilla (11–48 orifices; median, 27 orifices)[16] presents a challenge for recannulation of a specific duct. This may have been compounded by our study design, which permitted cannulation of duct orifices that were not producing NAF, because it is likely that NAF production, in conjunction with location information recorded on a grid, provides valuable visual cues for reidentifying a specific duct orifice. The insufficient sample rate on relavage for ducts that initially returned atypical cells was 12 of 42 samples (29%), essentially identical to the insufficient sample rate for the initial series of lavages. Possible reasons for the failure to obtain an adequate sample from ducts initially yielding atypical cells include inadvertent cannulation of a different duct orifice, regression of an atypical proliferative lesion either as a consequence of the initial lavage procedure or for reasons

unrelated to the initial lavage, and ductal injury at the time of the initial lavage that precluded satisfactory recannulation and lavage. The same operator performed all of the lavage procedures, and insufficient sample rates were stable over time (Fig. 1), excluding differences in operator experience or technique as factors in the assessment of reproducibility. It has been suggested that the use of normal saline for NDL can induce artifactual atypia and that plasmolyte is a superior lavage solution. Because all initial and repeat lavages were performed using normal saline, this is unlikely to have influenced reproducibility rates.

Finally, interobserver and intraobserver variation in cytology scoring can impact the reproducibility of cytologic screening tests. Interobserver variability was excluded by having the same cytopathologist evaluate all of the samples from this study. It is possible that atypia was over-called in the earliest period of the study, resulting in lower atypia rates for the repeat lavages, but this is unlikely, because the frequency of diagnosis of mild or marked atypia for the initial lavages was stable over time (Fig. 1).

It is likely that a combination of physiologic and technical factors accounted for the low reproducibility of repeat lavage measured in this series. Neither marked atypia on the initial lavage nor a 5-year Gail risk $\geq$ 1.7% predicted atypia on repeat lavage, but there was a trend toward improved reproducibility when the atypia initially was diagnosed in a NAF-producing duct. We previously reported that the frequency of lavage atypia is similar for patients with a 5-year Gail risk $<$ 1.7% and $\geq$ 1.7% and for ducts that produce NAF compared with ducts that do not produce NAF.[17] The high prevalence of lavage atypia noted in this and prior studies, combined with a low reproducibility, makes it unlikely that a single NDL demonstrating either mildly or markedly atypical cells will predict a high risk for breast carcinoma. Adjunctive tests, such as tumor suppressor gene methylation status,[18] loss of heterozygosity analysis,[19] or chromosome copy number determination,[20,21] are feasible for NDL samples and may improve the predictive value of NDL cytology.

Lavage atypia was associated with significant MRI abnormalities in only 1 of 24 breasts. The atypia in this breast was classified as marked and ultimately was diagnosed as ductal carcinoma in situ. In 23 breasts with atypical lavage results, there were no reproducible MRI findings. This is in marked contrast to a recent series that identified MRI abnormalities in six of seven breasts with atypical lavages.[11] Only one of those breasts was biopsied, and the diagnosis was atypical ductal hyperplasia. We performed MRI only in women whose third-party payors agreed to reimburse for the test (17 of 32 women). This is likely to have biased our results; however, given the low frequency of MRI findings, we currently are performing MRI only if marked atypia is confirmed on repeat lavage.

Ductal lavage is an excellent tool for retrieving breast epithelial cells, but the reproducibility of serial sampling is poor. In addition, lavage atypia is associated only infrequently with MRI findings but may represent mammographically occult DCIS. Until prospective studies validate lavage atypia as a marker for breast carcinoma risk, it is best to use it in the context of clinical trials.

REFERENCES

1. Wrensch MR, Petrakis NL, King EB, et al. Breast cancer incidence in women with abnormal cytology in nipple aspirates of breast fluid. *Am J Epidemiol.* 1992;135:130–141.
2. Wrensch MR, Petrakis NL, Miike R, et al. Breast cancer risk in women with abnormal cytology in nipple aspirates of breast fluid. *J Natl Cancer Inst.* 2001;93:1791–1798.
3. Fabian CJ, Kimler BF, Zalles CM, et al. Short-term breast cancer prediction by random periareolar fine-needle aspiration cytology and the Gail risk model. *J Natl Cancer Inst.* 2000;92:1217–1227.
4. Dupont WD, Parl FF, Hartman WH, et al. Breast Cancer risk associated with proliferative breast disease and atypical hyperplasia. *Cancer.* 1993;71:1258–1265.
5. Fisher B, Constantino JP, Wickerham DL, et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst.* 1998;90:1371–1388.
6. Port ER, Montgomery LL, Heerdt AS, Borgen PI. Patient reluctance toward tamoxifen use for breast cancer primary prevention. *Ann Surg Oncol.* 2001;8:580–585.
7. Morrow M, Vogel V, Ljung BM, O'Shaughnessy JA. Evaluation and management of the woman with an abnormal ductal lavage. *J Am Coll Surg.* 2002;194:648–656.
8. O'Shaughnessy JA, Ljung BM, Dooley WC, et al. Ductal lavage and the clinical management of women at high risk for breast carcinoma. *Cancer.* 2002;94:292–298.
9. Dooley WC, Ljung B, Veronesi U, et al. Ductal lavage for the detection of cellular atypia in women at high risk for breast cancer. *J Natl Cancer Inst.* 2001;93:1624–1632.
10. Brogi E, Robson M, Panageas KS, Casadio C, Ljung B-M, Montgomery L. Ductal lavage in patients undergoing mastectomy for mammary carcinoma. A correlative study. *Cancer.* 2003;98:2170–2176.
11. Hartman A-R, Daniel BL, Kurian AW, et al. Breast magnetic resonance image screening and ductal lavage in women at high genetic risk for breast carcinoma. *Cancer.* 2004;100:479–489.
12. Zera RT, Danielson D, Van Camp JM, et al. Atypical hyperplasia, proliferative fibrocystic change, and exogenous hormone use. *Surgery.* 2001;130:732–737.
13. Tavassoli FA. The influence of endogenous and exogenous reproductive hormones on the mammary glands with emphasis on experimental studies in rhesus monkeys. *Verh Dtsch Ges Pathol.* 1997;81:514–520.

14. Harvell DM, Strecker TE, Tochacek M, et al. Rat strain-specific actions of 17beta-estradiol in the mammary gland: correlation between estrogen-induced lobuloalveolar hyperplasia and susceptibility to estrogen-induced mammary cancers. *Proc Natl Acad Sci USA.* 2000;97:2779–2784.
15. Mitchell G, Trott PA, Morris L, Coleman N, Sauter E, Eeles RA. Cellular characteristics of nipple aspiration fluid during the menstrual cycle in healthy premenopausal women. *Cytopathology.* 2001;12:184–196.
16. Going JJ, Moffat DF. Escaping from Flatland: clinical and biological aspects of human mammary duct anatomy in three dimensions. *J Pathol.* 2004;203:538–544.
17. Johnson-Maddux A, Ashfaq R, Naftalis EZ, Leitch AM, Hoover S, Euhus DM. Patient and duct selection for nipple duct lavage. *Am J Surg.* 2004;188:390–394.
18. Evron E, Dooley WC, Umbricht CB, et al. Detection of breast cancer cells in ductal lavage fluid by methylation-specific PCR. *Lancet.* 2001;357:1335–1336.
19. Isaacs C, Cavalli LR, Cohen Y, et al. Detection of LOH and mitochondrial DNA alterations in ductal lavage and nipple aspirate fluids from high-risk patients. *Breast Cancer Res Treat.* 2004;84:99–105.
20. Yamamoto D, Senzaki H, Nakagawa H, Okugawa H, Gondo H, Tanaka K. Detection of chromosomal aneusomy by fluorescence in situ hybridization for patients with nipple discharge. *Cancer.* 2003;97:690–694.
21. King BL, Tsai SC, Gryga ME, et al. Detection of chromosomal instability in paired breast surgery and ductal lavage specimens by interphase fluorescence in situ hybridization. *Clin Cancer Res.* 2003;9:1509–1516.